DOI: https://doi.org/10.52756/ijerr.2026.v50.009 *Int. J. Exp. Res. Rev.*, Vol. 50: 110-118 (2026)

Original Article | *Peer Reviewed* | *Open Access*

International Journal of Experimental Research and Review (IJERR)
© Copyright by International Academic Publishing House (IAPH)
ISSN: 2455-4855 (Online)
www.iaph.in
ISSN 2455-4855

Hierarchical Ensemble Learning with Feature-Specific Classifiers for Heart Sound Classification

Aruna P. Kharat[1]* and Vipulsangram K. Kadam[2]

[1]Dr. Babasaheb Ambedkar Marathwada University, Aurangabad- 431004, Maharashtra, India;
[2]Department of Electronics & Telecommunication, P.E.S. College of Engineering, Aurangabad- 431004, Maharashtra, India

E-mail/Orcid Id:

AK, arunakharat62@gmail.com, https://orcid.org/0000-0002-7074-991X; *VK,* vipulsangram@pescoe.ac.in, https://orcid.org/0000-0001-7270-8225

Article History:

Received: 17th Nov., 2025
Accepted: 24th Mar., 2026
Published: 30th Apr., 2026

Keywords:
Heart Sound Classification, Phonocardiogram (PCG), Ensemble Machine Learning, Feature Extraction, Non-Invasive Diagnosis

How to cite this Article:
Aruna P. Kharat and Vipulsangram K. Kadam (2026). Hierarchical Ensemble Learning with Feature-Specific Classifiers for Heart Sound Classification. *International Journal of Experimental Research and Review, 50*, 110-118.
DOI:
https://doi.org/10.52756/ijerr.2026.v50.009

Abstract: Heart sound is one of the primary and most significant screening cardiac function screening techniques because it is non-invasive and requires fewer resources, especially in health facilities with limited resources. Nonetheless, manual evaluation is subjective and relies on clinicians' skills, potentially leading to variability and misdiagnosis. The proposal in this study is a machine learning ensemble design that will help automate the classification of phonocardiogram (PCG) signals into normal and abnormal categories. The proposed system combines time-domain, frequency-domain and time–frequency characteristics to store complete heart-sound characteristics. Three classifiers, random forest, Support Vector Machine (SVM) and gradient boosting, are used as base learners and aggregated using weighted voting to improve classification performance. Pre-processing, such as filtering, normalization and segmentation, is performed to improve signal quality and ensure accurate feature extraction. Experimental data show that the proposed model achieves 97.5% accuracy, 97.9% sensitivity and 96.7% specificity, outperforms more traditional single classifiers, and is equally effective as more complex deep learning models while incurring a lower computational cost and offering higher interpretability. The offered method is economical, dependable, and effective in the early detection of heart defects and, as such, is most applicable in primary care, rural settings and mobile diagnostic tools.

Introduction

Heart diseases have remained among the major causes of death in the world today and once detected early and properly diagnosed in their initial stages, it becomes vital in mitigating the risks involved and enhancing patient outcomes. Cardiac auscultation is considered one of the most basic, non-invasive, and cost-efficient methods of assessing cardiac viability among other diagnostic methods. It is especially useful in primary care and resource-limited settings where no elaborate diagnostic modalities, such as echocardiography and cardiac imaging, are available. Although widely used, traditional auscultation is highly subjective and largely reliant on the clinician's experience. The detection of small anomalies, such as murmurs, may be difficult, particularly when there is environmental noise, physiological variability and the superimposition of acoustic patterns. Moreover, auscultation results are difficult to record and analyze retrospectively, which limits their use in longitudinal monitoring and mass screening. The recent innovations in signal processing and machine learning have made it possible to create automated systems of phonocardiogram (PCG) signal analysis, i.e., the digital representation of sounds produced by the heart. Such systems offer the opportunity to improve diagnostic accuracy by deriving meaningful features from heart sound signals and classifying them as normal or abnormal (Khan et al., 2020; Aziz et al., 2020). Nevertheless, PCG signals are nonstationary and complex and are usually contaminated by noise and motion artifacts, making them extremely difficult to analyze reliably (Chowdhury et al., 2020). Conventional machine learning methods have used

*Corresponding Author: arunakharat62@gmail.com

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

handcrafted features with Support Vector Machines, k-Nearest Neighbor, and Artificial Neural Networks as classifiers, achieving moderate performance (Zheng et al., 2022; Rath et al., 2022). More recently, deep learning models have shown better predictive performance through learning features from raw data and their ability to learn and predict in low-resource settings is limited by their high computational demands and uninterpretable nature (Oh et al., 2020; Mahmood et al., 2024). To overcome these constraints, this paper introduces a hierarchical ensemble-based heart sound classification method. It has been demonstrated that ensemble learning methods enhance classification performance by integrating multiple models and leveraging complementary feature representations (Baydoun et al., 2020; Chen et al., 2022). In the proposed framework, several feature domains are incorporated -time, frequency, and time-frequency- to provide the complementary properties of PCG signals. The use of feature-specific classifiers in the form of Random Forest, Support Vector Machine and Gradient Boosting is used and the predictions are integrated using a weighted voting mechanism. This design has high classification accuracy, interpretability, and computational efficiency. The suggested solution is to offer a high-quality, scalable and economical solution for detecting cardiac anomalies at an early stage. It can be especially used in primary healthcare settings, rural settings and mobile diagnostics, thus enhancing access and equity in cardiac care.

Literature Review

Automatic classification of heart sound signals has developed considerably over the last 20 years, driven by advances in digital signal processing, the availability of publicly available signals, and the growing need to provide access to cardiac screening services. Studies in this field have already gone beyond the classic feature-based models to more sophisticated machine learning and deep learning models and ensembles, which have improved diagnostic performance (Chen et al., 2022; Nabhan Homsi et al., 2017).

Evolution of Heart Sound Analysis

The invention of the phonocardiogram (PCG) marked the beginning of automated analysis of the heart's acoustic signals, enabling visualization and quantitative analysis. Initial studies were more concerned with signal segmentation and the important phases of the heart, which included S1, systole, S2, and diastole. As signal processing methods advanced, scholars sought to extract salient features from PCG signals and investigate time-domain, frequency-domain, and time–frequency representations. The understanding of the non-stationary character of heart sounds led to the implementation of wavelet-based approaches, which offered better time and frequency resolution and formed the basis for machine learning (Chowdhury et al., 2020; Aldhahab et al., 2024).

Feature Extraction Methods

In heart sound classification, feature extraction is important because the model's discriminative ability depends on it. Time-domain features describe the statistical and temporal properties of the signal, such as mean amplitude, variance, skewness, kurtosis and energy distribution. Systolic and diastolic durations are also some of the temporal intervals that give essential diagnostic data. In addition to those measures, Shannon entropy can be used to measure signal complexity and irregularity (Khan et al., 2020).

Frequency-domain features provide clues to the spectral makeup of heart sounds, which are useful for identifying murmurs. Widespread ones are spectral centroid, bandwidth, and spectral flatness. Mel-frequency cepstral coefficients (MFCCs), widely used in speech processing, have been shown to be highly effective for characterizing heart sounds, as they can mimic human auditory perception (Rahmani et al., 2024; Abduh et al., 2020).

The non-stationary characteristics of PCG signals are met by time-frequency analysis. Time-frequency analysis methods, including the short-time Fourier transform and the discrete wavelet transform, enable simultaneous analysis of temporal and spectral variations. The 1-dimensional wavelet-based features, such as scale-wise energy and entropy, have shown good results in recognizing pathological patterns (Lee et al., 2023; Xiang et al., 2023).

Classical Machine Learning Solutions.

Conventional machine learning techniques have been widely used for classifying heart sounds using handcrafted features. Support Vector Machines (SVMs) have gained widespread popularity because they can classify high-dimensional data and model non-linear decision boundaries, but they are prone to feature scaling and scale poorly with large data sets. k-Nearest Neighbors (k-NN) is an easy-to-understand, but simple, method of classification, because it can classify high-dimensional data, and has a non-linear decision boundary, albeit with a very high scaling cost. Random Forest classifiers are robust and provide insights into feature importance by using ensemble decision trees, whereas Artificial Neural Networks (ANNs) can learn nonlinear relationships but require tuning to prevent overfitting.

DOI: https://doi.org/10.52756/ijerr.2026.v50.009

These methods also achieve moderate performance, with reported accuracies ranging from 85 to 92% (Zheng et al., 2022; Rath et al., 2022).

Deep Learning Approaches

With the development of deep learning, heart sound classification has greatly improved, as it enables end-to-end learning from raw signals or spectrograms. Convolutional Neural Networks (CNNs) are especially useful for extracting spatial information from time-frequency representations, whereas Recurrent Neural Networks (RNNs) are used to capture temporal dependencies in sequence data. More recently, CNNs combined with attention mechanisms or transformer architectures have reached state-of-the-art performance (Noman et al., 2019; Jamil et al., 2023; Hosseinzadeh et al., 2024).

Although they have achieved accuracies of more than 95%, deep learning models have been associated with high data requirements (large datasets), high computational cost, and limited interpretability. These constraints hinder their use in practical clinical settings, particularly in settings with limited resources (Oh et al., 2020; Mahmood et al., 2024).

Learning Ensemble Methodologies.

It was also discovered that ensemble learning is a potent approach to enhancing classification performance by combining multiple models. The complementary properties of various classifiers are used in techniques such as bagging, boosting and voting to reduce variance and enhance robustness. Ensemble techniques have also shown greater resistance to noise and variability in heart sound classification, achieving higher diagnostic accuracy (Potes et al., 2016; Humayun et al., 2018; Nogueira et al., 2019; Baydoun et al., 2020; Sarker et al., 2025).

But many current ensemble methods use homogeneous feature encodings or direct feature concatenation, without making extensive use of complementary features across different feature spaces. The practical usefulness of complex ensemble architectures can also be limited by the computational overhead they can create (Singh et al., 2023; Rahman et al., 2025).

Datasets and Benchmarking

The availability of publicly accessible datasets has been instrumental to research in this area. PhysioNet/CinC Challenge 2016 is a widely used benchmark dataset comprising 2,435 PCG recordings from 1,297 subjects, including both healthy and abnormal heart sounds. This dataset is representative of variability in recording conditions, noise levels, and patient demographics, which are related to the real world, and thus it can be used to assess model robustness (Potes et al., 2016).

Gaps and Motivation in Research.

Although much has been done, many challenges remain. The available literature typically focuses on individual feature domains or on simple strategies for feature fusion, which do not enable exploration of complementary signal properties. Deep learning models are highly accurate but not interpretable and consume substantial computational resources. Moreover, data imbalance, signal fluctuations, and reduced attention to demographic variables continue to influence model generalization (Yakut et al., 2022).

In a bid to overcome such shortcomings, this paper suggests a hierarchical ensemble framework that combines time-domain, frequency-domain and time-frequency attributes using feature-specific classifiers. By integrating the performance of Random Forest, Support Vector Machine, and Gradient Boosting using a weighted voting system, the proposed methodology is expected to perform well, enhance robustness and improve computational efficiency, making it suitable for practical healthcare settings.

Methodology

The suggested heart sound classification framework is considered a pipeline that will provide both diagnostic predictions (proposed as accurate), as well as computational efficiency and interpretability, and will be structured and systematic, as phonocardiogram (PCG) recording will be converted into the correct diagnostic prediction. The methodology comprises six main steps: dataset selection, signal pre-processing, cardiac cycle segmentation, feature extraction, base-learner training, and decision fusion via ensembles. Each of these steps is designed to address the nature of PCG analysis problems, noise contamination, signal variation, and class bias (Chowdhury et al., 2020; Khan et al., 2020; Tariq et al., 2022).

Dataset Description

The analysis will use PhysioNet/CinC Challenge 2016 data, which is a standard platform for heart sound classification. The sample comprises 2,435 PCG recordings from 1,297 subjects across various clinical settings, with varying patient demographics, recording conditions, and noise levels (Potes et al., 2016). Of these recordings, 665 are deemed abnormal and depict various cardiac pathologies, such as murmurs and gallop rhythms, while the other recordings are considered normal. The dataset is particularly well-suited for

measuring the strength and practicality of classification models due to its heterogeneity.

Signal Pre-processing

Pre-processing is important for improving signal quality and enabling reliable feature extraction. All recordings are initially resampled to a constant 1000 Hz to ensure consistency across the dataset. Ahead and behind silence blocks are eliminated to cut out irrelevant signals. A zero-phase band-pass filter with cutoff frequencies of 25 Hz and 400 Hz is used to reduce artifacts from low-frequency motion and high-frequency noise while preserving other components of the clinically significant heart sound (Chowdhury et al., 2020). The model of the filtering process is:

$$y[n] = \sum_{k=0}^{M} b_k\, x[n-k] - \sum_{j=1}^{N} a_j\, y[n-j]$$

where b_k and a_j represent feed-forward and feedback coefficients, respectively.

The filtered signal is normalized to its maximum amplitude to minimise inter-recording variations. Also, sensor motion leads to impulsive noise, which is reduced using median filtering and outlier suppression algorithms (Aziz et al., 2020).

Cardiac Cycle Segmentation

To perform any meaningful analysis, continuous PCG signals must be segmented into individual cardiac cycles. The proposed structure is a hybrid one that uses probabilistic modeling and signal-based techniques. The Hidden Semi-Markov Model (HSMM) estimates the probability of each sample state using a logistic regression model. This method explicitly represents state durations, allowing sufficient fragmentation into physiological phases: S1, systole; S2, diastole (Potes et al., 2016). The probability function is given as:

$$p(s = 1 \mid z) = \frac{1}{1 + \exp[-(\beta_0 + \beta^T z)]}$$

An envelope-based Hilbert transform approach is also used to increase robustness. The analytic signal is calculated and the magnitude is used to identify peaks that are related to heart sound events. Cardiac cycles are determined by successive peaks, and only cycles in the range of 0.4 to 1.2 seconds are stored to ensure physiological validity (Arslan et al., 2022).

Feature Extraction

Feature extraction transforms segmented cardiac cycles into compact, discriminative representations. To represent the non-stationary and complex characteristics of heart sounds, three domains are considered complementary, i.e., features are extracted across these domains. Time-domain statistics have been proposed as features that describe amplitude distribution and time behaviour (mean, variance, skewness, kurtosis and signal energy). The discrete Fourier transform is used to compute frequency-domain features, such as spectral centroid, bandwidth and spectral flatness, which provide information about the frequency distribution. Mel-frequency cepstral coefficients (MFCC) and discrete

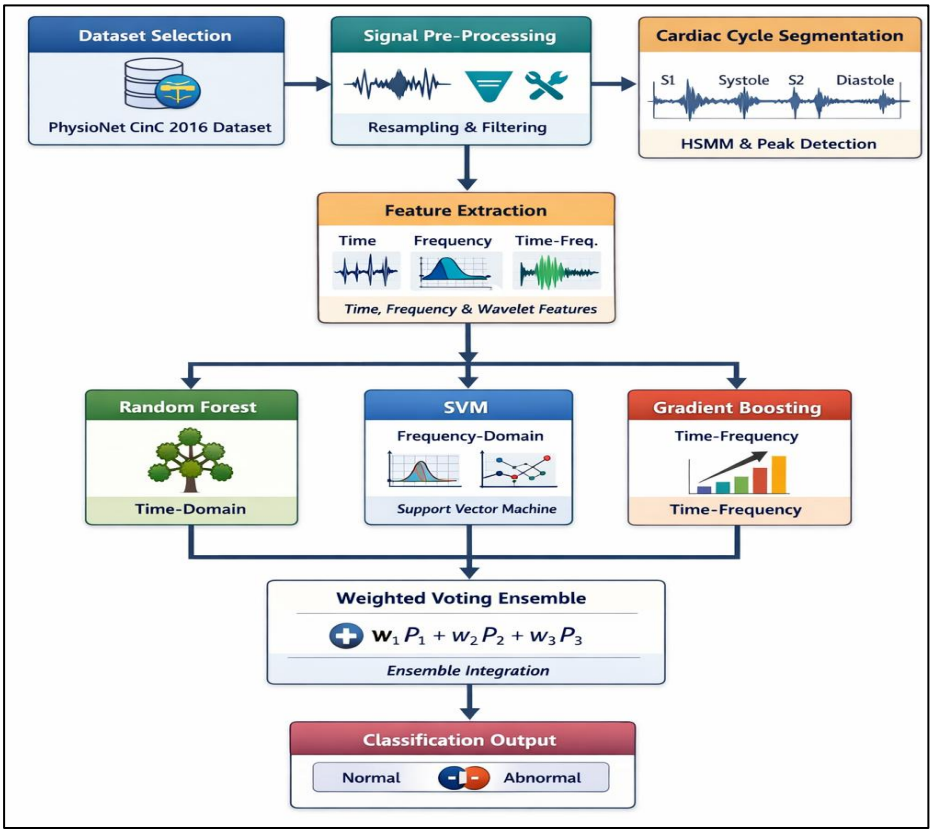

Figure 1. Flowchart diagram of methodology.

DOI: https://doi.org/10.52756/ijerr.2026.v50.009

wavelet transform are used to extract time-frequency features. MFCCs provide perceptually meaningful spectral representations (Rahmani et al., 2024), whereas wavelet-based features capture local changes in signal energy across scales (Lee et al., 2023). The combination of feature domains guarantees full characterization of heart sound signals.

Base Learner Training

Three independent classifiers are trained on different feature sets to exploit the merits of these representations. The time-domain features are processed using a Random Forest classifier, which offers strength and the ability to capture interactions among complex features (Zheng et al., 2022). A Support Vector Machine (SVM) with a radial basis function (RBF) kernel is trained on frequency-domain features and is capable of capturing non-linear decision boundaries. Time-frequency features are trained in Gradient Boosting classifier, which enhances its performance by refining the results with misclassified instances (Baydoun et al., 2020). Probabilistic outputs from each classifier are integrated using an ensemble combination.

Ensemble-Based Decision Fusion

The last category is determined by a weighted-voting ensemble mechanism that combines the predictions of the individual classifiers based on their performance on the validation set. Ensemble learning has been shown to enhance classification accuracy through the synergistic advantages of individual models (Chen et al., 2022; Nabhan Homsi et al., 2017). A combination of the classes' probability is a sum of the individual classifier probabilities formulated as a weighted probability, and the ultimate class is assigned with the help of the maximum probability criterion. This is a good strategy that incorporates complementary feature information, thereby enhancing classification accuracy, robustness and generalization.

The suggested methodology will combine sophisticated signal processing and machine learning approaches in a single framework. The system attains a balance between interpretability, high accuracy and computational efficiency by combining feature-specific classifiers using a weighted ensemble strategy. This is why it is especially applicable to the implementation in the primary healthcare and resource-limited settings, where the availability and high-quality diagnostic tools cannot be ignored.

Results and Discussion

This part presents the experimental findings of the proposed hierarchical ensemble model for classifying heart sounds and discusses them in detail. The proposed model is compared against baseline classifiers, and state-of-the-art methods and the effect of single components on the performance of the suggested model is thoroughly analyzed, as well as the implication on clinical implementation. The findings show that the suggested framework has better performance and is still interpretable and computationally efficient.

Experimental Setup

The experiments have been done on a standardized computational environment, which guarantees reproducibility and fair comparability of the models. Stratified sampling was used to divide the dataset so that the distribution of classes between the training, validation and the test groups would be maintained. Particularly, 60 percent of the data was used to train, 20 percent to validate and 20 percent to test. The validation set was only utilized in hyper parameter tuning and computing weights of the ensemble, and the test set was not viewed at all when developing the model to allow an unbiased estimation of the generalization performance. The Random Forest classifier used had 100 decision trees and the maximum depth was not limited to enable the model to find out more complicated relations in the time-domain features. The classifier used was the SVM, with a radial basis function kernel; the regularization parameter and kernel width were optimized via grid search on the validation set. A gradient Boosting classifier was set to 100 boosting stages, and a learning rate of 0.1, but tree depth was restricted to three to avoid overfitting and large enough to allow the interactions of wavelet features. Any performance measures presented are based on the held-out test set; hence, the results reflect the models' performance on unseen data rather than on training examples.

Overall Classification Performance

Table 1 shows the general performance of the hierarchical ensemble framework proposed versus the single classifier and the best classifier in the literature. The proposed model has 97.5% accuracy, 97.9% sensitivity, 96.7% specificity and F1-score of 98.0, which is a significant improvement over the traditional approach.

The proposed model improves accuracy by approximately 5–6% over baseline classifiers, confirming the effectiveness of ensemble learning in leveraging complementary feature representations (Chen et al., 2022; Baydoun et al., 2020).

The model achieves a balanced performance, with sensitivity (97.9%) and specificity (96.7%) both exceeding 96%, which is critical for clinical screening

DOI: https://doi.org/10.52756/ijerr.2026.v50.009

Table 1. Performance Comparison of Proposed Method Against Baseline and Existing Models.

Model	Type	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
Hierarchical Ensemble (Proposed)	Feature-specific ensemble	97.5	97.9	96.7	98.0
SVM (single classifier)	Baseline model	91.95	92.78	91.14	-
k-NN (k = 3)	Baseline model	91.9	91.41	92.48	-
PCTMF-Net	Existing model (2023)	93.0	~92	~91	~92
CTENN (Conv + Transformer)	Existing model (2023)	95.7	-	-	-

applications.

Comparison with Deep Learning and Complex Ensembles

Table 2 extends the comparison to include recent deep learning models and complex ensemble architectures, providing context for the proposed model's performance within the broader landscape of heart sound classification research.

The proposed model achieves performance comparable to complex ensemble architectures such as ENACT Heart, while maintaining significantly lower computational complexity. Unlike deep learning approaches, which require large datasets and high computational resources, the proposed model offers efficient and interpretable performance (Oh et al., 2020; Mahmood et al., 2024).

Contribution of Individual Feature Groups

A critical aspect of the proposed framework is its use of feature-specific base learners, each trained on a distinct feature representation. Table 3 presents the performance of individual classifiers trained on each feature group, alongside the final ensemble performance, to demonstrate the benefit of combining complementary features.

Each feature domain contributes unique discriminative information. Wavelet-based features show the highest individual performance, while ensemble integration improves accuracy by 2–3%, validating the complementary nature of multi-domain features (Rahmani et al., 2024; Lee et al., 2023).

Weighted Voting Analysis

The ensemble assigns weights based on validation accuracy:

1. Random Forest: 0.32
2. SVM: 0.33
3. Gradient Boosting: 0.35

This balanced weighting ensures that each classifier contributes proportionally, improving overall robustness and reducing bias (Nabhan Homsi et al., 2017).

Confusion Matrix Analysis

The model correctly classifies:

: 97.8% of normal cases

: 97.6% of abnormal cases

Misclassifications are mainly due to:

: Low-intensity murmurs (false negatives)

: Noise artifacts (false positives)

This indicates the importance of improved pre-processing techniques (Chowdhury et al., 2020).

Computational Efficiency

Maintaining computer efficiency to enable implementation in resource-restrained environments was among the design goals of the proposed framework. A comparison with the proposed model was conducted in terms of training time, inference time, and memory footprint. The entire ensemble took about 12 minutes to train on an average workstation with 16 GB of RAM and 2.8 GHz of processing power. Inference-classification of an individual PCG recording took an average of 0.15 seconds, far less than the time ranges used in real-time or near-real-time clinical practice. The overall memory consumption of the trained ensemble, including the three underlying base learners, was about 45 MB, and thus can be readily implemented on small chip-based systems, home mobile, or small diagnostic systems. These computational properties are contrasted with deep learning methods that typically require graphics processing units for computation and can be memory-intensive, with memory footprints in the Megabyte range.

Table 2. Performance Comparison with Deep Learning and Complex Ensemble Models.

Model	Accuracy (%)	Precision (%)	Recall (%)
Hierarchical Ensemble (Proposed)	97.5	98.1	97.9
ENACT Heart (Mixture-of-Experts)	97.52	-	-
PCTMF-Net	93.0	~92	~91
CTENN (Conv-Transformer)	95.7	-	-

DOI: https://doi.org/10.52756/ijerr.2026.v50.009

The efficiency of the suggested model is achieved without compromising performance, as evidenced by the accuracy measurements above.

Comparison with Previous Studies

The framework's performance is comparable to prior research on heart sound classification. Previous studies report:

1. Traditional ML accuracy: 85–92%
2. Deep learning precision: 9397% (Chen et al., 2022)

The provided model has an accuracy of 97.5%, making it one of the best-performing methods in terms of interpretability and efficiency.

Clinical Implications

The suggested model has high clinical relevance, achieving high sensitivity (97.9%) and specificity (96.7%) and reliably detecting abnormal heart sounds, with minimal false alarms. This equilibrium is essential when performing screening, especially in resource-limited environments.

Its low computational requirements can be integrated into portable devices such as electronic stethoscopes, enabling real-time analysis in primary care and rural settings. Besides, interpretable features and transparent classifiers help increase clinical trust in black-box models (Mahmood et al., 2024).

In general, the framework is an effective and scalable solution to enhance access to early cardiac diagnosis (Chen et al., 2022).

Limitations

Although these have promising outcomes, there are a number of limitations. The model is tested on a single dataset, which may constrain generalization across a wide range of clinical contexts (Potes et al., 2016). The sample is not demographically diverse, either, which may affect the ability to assess performance across various populations.

The use of handcrafted features might limit the model's ability to capture complex patterns, suggesting the need to implement hybrid solutions that combine traditional and deep learning techniques (Rahmani et al., 2024). The system also performs only binary classification, so it cannot detect specific cardiac conditions.

Summary

The suggested ensemble model achieves a high accuracy of 97.5%, outperforming traditional models and matching the state-of-the-art. It is equipped with complementary features and classifiers that provide robustness, efficiency, and interpretability (Baydoun et al., 2020).

Such features render it appropriate for use in primary care and low-resource settings in real-world settings.

Table 3. Performance of Base Learners vs. Final Hierarchical Ensemble.

Feature Set / Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
Time-domain features → Random Forest	94.2	94.8	93.5
Frequency-domain features → SVM	94.5	93.7	94.9
Time-frequency (wavelet) features → Gradient Boosting	95.3	95.0	94.7
Hierarchical Ensemble (Weighted Voting)	97.5	97.9	96.7

Conclusion and Future Scope

Conclusion

This paper proposes a hierarchical ensemble system for heart sound classification using phonocardiogram (PCG) signals. The method combines time-domain, frequency-domain and time-frequency characteristics with feature-specific classifiers-Random Forest, Support Vector machine and Gradient Boosting-weighted voting scheme.

The model has a high accuracy of 97.5%, a sensitivity of 97.9%, a specificity of 96.7% and an F1-score of 98.0, providing better results than traditional classifiers and comparable results to advanced deep learning approaches. The merging of complementary properties makes it more robust and generalized in a wide variety of situations.

Moreover, the framework has limited computational complexity and interpretability and is therefore applicable to primary healthcare, rural regions and portable diagnostics. In general, the paper indicates that ensemble learning is an effective and viable method of automated cardiac screening.

Future Scope

Future research can focus on developing a multi-class classification model to detect specific cardiac disorders. Combining deep learning and an ensemble could also improve performance.

Generalization and fairness will be enhanced by using

DOI: https://doi.org/10.52756/ijerr.2026.v50.009

more and varied datasets. Latency Real-time embedded systems can be developed to be deployed in mobile health platforms and portable devices.

Continuous monitoring and remote diagnosis can be achieved through further enhancements to noise-reduction methods and by integrating them with telemedicine and IoT-based systems.

Conflict of interest

None

References

Abduh, Z., Nehary, E. A., Abdel Wahed, M., & Kadah, Y. M. (2020). Classification of heart sounds using fractional Fourier transform-based mel-frequency spectral coefficients and traditional classifiers. *Biomedical Signal Processing and Control, 57*, 101788. https://doi.org/10.1016/j.bspc.2019.101788

Aldhahab, A. Q., Ismael, M. R., & Abd, H. J. (2024). Classification of Phonocardiogram Signals Using the Wavelet Scattering Transform and Equilibrium Optimization Approach. *IEEE Access, 1*. https://doi.org/10.1109/access.2024.3518635

Arslan, Ö., & Karhan, M. (2022). Effect of Hilbert-Huang transform on classification of PCG signals using machine learning. *Journal of King Saud University - Computer and Information Sciences, 34*(10), 9915–9925. https://doi.org/10.1016/j.jksuci.2021.12.019

Aziz, S., Khan, M. U., Alhaisoni, M., Akram, T., & Altaf, M. (2020). Phonocardiogram Signal Processing for Automatic Diagnosis of Congenital Heart Disorders through Fusion of Temporal and Cepstral Features. *Sensors, 20*(13), 3790. https://doi.org/10.3390/s20133790

Baydoun, M., Safatly, L., Ghaziri, H., & El Hajj, A. (2020). Analysis of heart sound anomalies using ensemble learning. *Biomedical Signal Processing and Control, 62*, 102019. https://doi.org/10.1016/j.bspc.2020.102019

Chen, J., Dang, X., & Li, M. (2022). Heart Sound Classification Method based on Ensemble Learning. *2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)*, 8–13. https://doi.org/10.1109/icsp54964.2022.9778383

Chowdhury, T. H., Poudel, K. N., & Hu, Y. (2020). Time-Frequency Analysis, Denoising, Compression, Segmentation, and Classification of PCG Signals. *IEEE Access, 8*, 160882–160890. https://doi.org/10.1109/access.2020.3020806

Hosseinzadeh, M., Haider, A., Malik, M. H., Adeli, M., Mzoughi, O., Gemeay, E., Mohammadi, M., Alinejad-Rokny, H., Khoshvaght, P., Porntaveetus, T., & Rahmani, A. M. (2024). Enhanced heart sound classification using Mel frequency cepstral coefficients and comparative analysis of single vs. ensemble classifier strategies. *Plos One, 19*(12), e0316645. https://doi.org/10.1371/journal.pone.0316645

Humayun, A. I., Khan, Md. T., Ghaffarzadegan, S., Feng, Z., & Hasan, T. (2018). An Ensemble of Transfer, Semi-supervised and Supervised Learning Methods for Pathological Heart Sound Classification. *Interspeech, 2018*, 127–131. https://doi.org/10.21437/interspeech.2018-2413

Jamil, S., & Roy, A. M. (2023). An efficient and robust Phonocardiography (PCG)-based Valvular Heart Diseases (VHD) detection framework using Vision Transformer (ViT). *Computers in Biology and Medicine, 158*, 106734. https://doi.org/10.1016/j.compbiomed.2023.106734

Khan, F. A., Abid, A., & Khan, M. S. (2020). Automatic heart sound classification from segmented/unsegmented phonocardiogram signals using time and frequency features. *Physiological Measurement, 41*(5), 55006. https://doi.org/10.1088/1361-6579/ab8770

Lee, J.A., & Kwak, K.C. (2023). Heart Sound Classification Using Wavelet Analysis Approaches and Ensemble of Deep Learning Models. *Applied Sciences, 13*(21), 11942. https://doi.org/10.3390/app132111942

Mahmood, A., Dhahri, H., Alhajlah, M., & Almaslukh, A. (2024). Enhanced Classification of Phonocardiograms Using Modified Deep Learning. *IEEE Access, 12*, 178909–178916. https://doi.org/10.1109/access.2024.3507920

Nabhan Homsi, M., & Warrick, P. (2017). Ensemble methods with outliers for phonocardiogram classification. *Physiological Measurement, 38*(8), 1631–1644. https://doi.org/10.1088/1361-6579/aa7982

Nogueira, D. M., Ferreira, C. A., Gomes, E. F., & Jorge, A. M. (2019). Classifying Heart Sounds Using Images of Motifs, MFCC and Temporal Features. *Journal of Medical Systems, 43*(6). https://doi.org/10.1007/s10916-019-1286-5

Noman, F., Ting, C.-M., Salleh, S.-H., & Ombao, H. (2019). Short-segment Heart Sound Classification Using an Ensemble of Deep Convolutional Neural Networks. *ICASSP 2019 - 2019 IEEE International*

Conference on Acoustics, Speech and Signal Processing (ICASSP), 1318–1322. https://doi.org/10.1109/icassp.2019.8682668

Oh, S. L., Jahmunah, V., Ooi, C. P., Tan, R.-S., Ciaccio, E. J., Yamakawa, T., Tanabe, M., Kobayashi, M., & Rajendra Acharya, U. (2020). Classification of heart sound signals using a novel deep WaveNet model. *Computer Methods and Programs in Biomedicine, 196*, 105604. https://doi.org/10.1016/j.cmpb.2020.105604

Potes, C., Parvaneh, S., Rahman, A., & Conroy, B. (2016). Ensemble of Feature: based and Deep learning: based Classifiers for Detection of Abnormal Heart Sounds. *2016 Computing in Cardiology Conference* (CinC), 43. https://doi.org/10.22489/cinc.2016.182-399

Rahman, M. T., & Shahnaz, C. (2025). LtXe-EnsNet: A Lightweight and Explainable Ensembled Deep Learning Model for Heart Sound Abnormality Classification from Sensor Data. *IEEE Sensors Letters, 9*(9), 1–4. https://doi.org/10.1109/lsens.2025.3596256

Rath, A., Mishra, D., Panda, G., & Pal, M. (2022). Development and assessment of machine learning based heart disease detection using imbalanced heart sound signal. *Biomedical Signal Processing and Control, 76*, 103730. https://doi.org/10.1016/j.bspc.2022.103730

Sarker, S., Fairuj Preotee, F., Muhammad, T., & Akhter, S. (2025). An Innovative Framework for Heart Sound Classification Integrating Adaptive Fuzzy Rank-Based Ensemble of Transfer Learning Models with Multi-Dimensional Features. https://doi.org/10.1101/2025.09.07.674773

Singh, A.S., Singh, A., & Singh, D. (2024). Ensemble Learning for Accurate Prediction of Heart Sounds using Gammatonegram Images. *Turkish Journal of Electrical Engineering and Computer Sciences, 32*(4), 555–573. https://doi.org/10.55730/1300-0632.4087

Singh, S. A., Devi, N. D., Singh, K. N., Thongam, K., D, B. R., & Majumder, S. (2023). An ensemble-based transfer learning model for predicting the imbalance heart sound signal using spectrogram images. *Multimedia Tools and Applications, 83*(13), 39923–39942. https://doi.org/10.1007/s11042-023-17186-9

Tariq, Z., Shah, S. K., & Lee, Y. (2022). Feature-Based Fusion Using CNN for Lung and Heart Sound Classification. *Sensors, 22*(4), 1521. https://doi.org/10.3390/s22041521

Xiang, M., Zang, J., Wang, J., Wang, H., Zhou, C., Bi, R., Zhang, Z., & Xue, C. (2023). Research of heart sound classification using two-dimensional features. *Biomedical Signal Processing and Control, 79*, 104190. https://doi.org/10.1016/j.bspc.2022.104190

Yakut, Ö., & Bolat, E. D. (2022). A high-performance arrhythmic heartbeat classification using ensemble learning method and PSD based feature extraction approach. *Biocybernetics and Biomedical Engineering, 42*(2), 667–680. https://doi.org/10.1016/j.bbe.2022.05.004

Zheng, Y., Guo, X., Wang, Y., Qin, J., & Lv, F. (2022). A multi-scale and multi-domain heart sound feature-based machine learning model for ACC/AHA heart failure stage classification. *Physiological Measurement, 43*(6), 65002. https://doi.org/10.1088/1361-6579/ac6d40

How to cite this Article:

Aruna P. Kharat and Vipulsangram K. Kadam (2026). Hierarchical Ensemble Learning with Feature-Specific Classifiers for Heart Sound Classification. *International Journal of Experimental Research and Review, 50*, 110-118.

DOI : https://doi.org/10.52756/ijerr.2026.v50.009

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.